OM nucleic · nucleic search, using sw model.

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

Run on: May   2, 2003, 14:33:45 ; Search time 83775 Seconds
                              (without alignments)
                              2.679 Million cell updates/sec

Title: us-09-673-476-1
Perfect score: 12732
Sequence: 1 acctgcgcttgcagagatca.............gtcggtgctggttggctcgc 12732

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 0.5

Searched: 2 seqs, 8815365 residues

Total number of hits satisfying chosen parameters: 4

Post-processing Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   1: *gbk:*
               1: /home/sdavid/sdavid-tmp/may03/chakrabarti476/AL123456.gbk:*
               2: /home/sdavid/sdavid-tmp/may03/chakrabarti476/AE005516.gbk:*

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

```
                  %       Query
Result   Score  Match  Length DB  ID                    Description
No.
  1    12732  100.0  4411529  1   AL123456    ACCESSION:AL123456
  2    2715.7   99.9  4403836  1   AE005516    ACCESSION:AE005516
c 3    3484.999   3.8 4403836  2   AE000516    ACCESSION:AE000516
c 4    483.599    3.8 4411529  1   AL123456    ACCESSION:AL123456
```

ALIGNMENTS

RESULT 1
AL123456
LOCUS   AL123456   4411529 bp   circular BCT   07-JUL-1998
DEFINITION Mycobacterium tuberculosis complete genome.
ACCESSION   AL123456
VERSION   AL123456
KEYWORDS   .
SOURCE   Mycobacterium tuberculosis.
ORGANISM   Mycobacterium tuberculosis
            Bacteria; Firmicutes; Actinomycetes; Mycobacterium;
            Mycobacteriaceae; Mycobacterium.
REFERENCE   1 (bases 1 to 4411529)
AUTHORS   Cole,S.T., Brosch,R., Parkhill,J., Garnier,T., Churcher,C.,
            Harris,D., Gordon,S.V., Eiglmeier,K., Gas,S., Barry III,C.E.,
            Tekaia,F., Badcock,K., Basham,D., Brown,D., Chillingworth,T.,
            Connor,R., Davies,R., Devlin,K., Feltwell,T., Gentles,S.,
            Hamlin,N., Holroyd,S., Hornsby,T., Jagels,K., Krogh,A., McLean,J.,
            Moule,S., Murphy,L., Oliver,S., Osborne,J., Quail,M.A.,
            Rajandream,M.A., Rogers,J., Rutter,S., Seeger,K., Skelton,S.,
            Squares,S., Squares,R., Sulston,J.E., Taylor,K., Whitehead,S. and
            Barrell,B.G.
TITLE   Deciphering the biology of Mycobacterium tuberculosis from the

JOURNAL   Nature 393 (6685), 537-544 (1998)
MEDLINE   98295987
REFERENCE   2 (bases 1 to 4411529)
AUTHORS   Parkhill,J.
TITLE   Direct Submission
JOURNAL   Submitted (11-JUN-1998) submitted on behalf of the Mycobacterium
            tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome
            Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique
            Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux,
            75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk
FEATURES   Location/Qualifiers
     source   1..4411529
                     /organism="Mycobacterium tuberculosis"
                     /db_xref="taxon:1773"
     gene   1..1524
                     /gene="dnaA"
                     /note="Rv0001"
                     /db_xref="EMBL:3261512"
     CDS   1..1524
                     /gene="dnaA"
                     /note="Rv0001, (MTV029.01), len: 507, dnaA, chromosomal
                     replication initiator protein, identical to TR:O69585
                     (EMBL:X92504) and highly similar to many eg. DNAA_ECOLI
                     P03004 chromosomal replication initiator protein (467 aa),
                     fasta scores: opt: 986 z-score: 1154.5 E(): 0, 43.2%
                     identity in 449 aa overlap. Contains PS00017
                     ATP/GTP-binding site motif A (P-loop) and PS00027 dnaA
                     protein signature. Note that the first base of this gene
                     has been taken as base 1 of the M. tuberculosis H37Rv
                     genomic sequence."
                     /codon_start=1
                     /transl_table=11
                     /product="dnaA"
                     /protein_id="CAA16238.1"
                     /db_xref="EMBL:3261512"
                     /translation="MIDDDGSSIIWNAVVSELNGEPKVDRVDSPSSDANLSAPLTPQQ
                     RAWLNLVQPLIVEGFALLSVPSSSFVQNEIERHLRAPICALSRRTGHQLICGVRTAP
                     PATDEADCITVPPSSENFATPSPDIITCNDEIDSSAAARCDQHSWSSYFPPSNICS
                     ATAGVTSLNRVTEDITPVTGASNRFAHAAALAIAEAPRAAYNFLHGESSLGK+HLL
                     HAAGNYAORLPGMRVKVVSTEEFNFINSLRDCRKVAFKKSRKVDVCLVCLLGFL
                     EGKEIIGEEFHIFNILINAKGIV:SSDRFPKQIATLEDRLRKFEWGLLIVDQPP
                     LETRIALKKKAQMRLAVPDVLELIASSERNIKEEGALIRVTAPASNKFFIDK
                     ALAEIVLRDLIAPANTMCISAAIXAATAEKFPTVEELPGP2KTRRLACSKQIAMYL
                     CREL:LVLSIPKIGVAFGEDHTIVMYAQRKILSEMAERKFEVFPDVKELT:NIRDRSKS"
     gene   2052..3260
                     /gene="dnaN"
                     /note="Rv0002"
                     /db_xref="EMBL:3261512"
     CDS   2052..3260
                     /gene="dnaN"
                     /note="Rv0002, (MTCY10H4.02), len: 402 aa, dnaN,
                     dna polymerase III beta chain, almost identical to O69590
                     (EMBL:X92504) dnaN (399 aa) (96.9% identity in 368 aa
                     overlap) and highly similar to many eg. DPOL_SUPCO P27203
                     dna polymerase III beta chain (eg 2) (376 aa)
                     fastascores: opt: 1393 z-score: 1567.3 E(): 0, 52.8%
                     identity in 397 aa overlap and overlaps and extends CDS i:
                     neighbouring cosmid M:CY10H4.01"
                     /codon_start=1
                     /transl_table=11
                     /product="dnaN"
                     /protein_id="CAA16233.1"
                     /db_xref="EMBL:3261513"
                     /translation="MDAATRVGLTDLKTRLRESFALAVSWVAKNLPARRAAVVLSS
                     VLLTGSDNSNISGFDVEVSAPAQVGAEIVSPGSVLVSGRLLSEIITRALENKFWNEV
                     EGNRVAITCGNARFSLPTMPVEYPLPILPEFTGEAELFAEAISQVKLNAGHDDI
                     LPMZCIRVELIGETVVLLAATDRFLAVREIKWSASSPGIRAAVLVPKILAAAKAS
                     IGSSLVRLSLGTRPGVSKRGLGISGNGKRSTRKFLHAEFPEVPKLPIEHIAVAVMI
                     VAEILIEAIKIVALVADRIAQVWMEFADGSVRLSAGADGNGRAEDLVVDVAGSPLTA
                     FNPTLTDGLSSLRSERVSGFTTACKVALLRPVSGDDRPVAGLNNGFPAVSUPYV
                     YLIMPVKLPG"
     gene   3280..4437

CDS

/gene="recF"
/note="Rv0003"
/db_xref="EMBL:3261645"
3280..4437
/gene="recF"
/note="Rv0003"
/product="recF"
/db_xref="GI:1552555"
/db_xref="SWISS-PROT:Q59556"
/translation="MYVRHLGLRDFRSWACVCLELHPGRTVFVGPNGYGKTNLIEALW
YSITLGSHRVSAILPLRVGTURAVIS..VVNDGREQAVDEEIATGRVNKARLNRSSV
RSTPDVVGVLRAVLPAPELGLIVRDPAQRRYEDLAIVRPPAIANAVAAFVEERVLPQ
HTACLKSVPGAARVRGDRGVVDFLEVWDSRLAEGSAELVAAKIDIVNQLAPEVKKAYQL
LAPERSAS...GYRASMDVTGPSEQSDIRFQL.AARILAALAARVRDEFAFELRVGCLCGPFH
RDGL.LRLQQPAKGFASHGCEAWSLAVALHLAAYQLLRVDGGEPVLLLDDVFAELGVM
RRRAIAIAARFSAEQV.WTAAVEEDIPAGWDARVHIDVRAG...TGSMSVVLLP"
/db_xref="SPTREMBL:P77573"
/codon_start=1
/transl_table=11

gene

4434..4997
/gene="Rv0004"
/note="Rv0004"
/db_xref="EMBL:3261645"
4434..4997
/gene="Rv0004"
/note="Rv0004"
/db_xref="GI:1552557"
/protein_id="CAB02425.1"
/product="hypothetical protein Rv0004"
/translation="MTGSVDRPVQNRGFERSMKSPGLDLVRRTLDEARAAARARGQDA3
RGRVASSVASGRVASRRHSWSGPCPLCRCPFQLCKKAARELAKKRGWSSVRVAEGWVLGGW
KITGPAAPSWRKIPHLMGRGPRDYYG"
(97.9% identity in 187 aa overlap)
/codon_start=1
/transl_table=11

gene

5123..7267
/gene="gyrB"
/note="Rv0005"
/db_xref="EMBL:3261645"
5123..7267
/gene="gyrB"
/note="Rv0005", (MTCY10H4.03), gyrB, dna gyrase subunit b,
len: 714, identical to GYRB_MYCTU P41514, contains PS00177
DNA topoisomerase II signature"
/db_xref="SWISS-PROT:P41514"
/db_xref="GI:1552558"
/protein_id="CAB02426.1"
/product="gyrB"
/translation="MGKNEAPRSALAPHGTVVDDFRRLXRHAIPEFSIR.VAAGK
KKAQDEVGAASIILLEGLEAVRKRKGWKIGSTFDRGLHHLIWEVVDNAVCEAMAGYAI
IVNVLEDGGVRVAVDGRG.PVAIHASSPTVDVVMTQLLHAGKPEDSPAYAISGGLH
GVGVSVVNAISPLFWSIKRDGYERSQVEKSEPQGLKQGAPTKKTGSTVRFWADPAV
HETTEYDPETVARRLQFMAFLNKGITINLTDERVQDEVDEVVSDVAEPKSASERA
AESTAPHKVASRDFHFEGGLTAVVNHHIINRTKNAIHSSIVDERSGKGGHEVEAMQWNA
GYSESVHTFAN..NTEGGTEHEEGFRSALTSVNKTAXRRKLLKUKDPNLTGDDIREG
LAAVISVKVSEPEFEGQTKTLGNTFVKSFPVQKVCNEQLIHWFEANPTDAKVVNKAV
SSAQARIARKAFELVHRKKA.DIGGLFGKLACRSTPRRSELIVVEGUSAGGSAKS
GRDSMFQALPIRGKIIVEKARIDRVIKNTEVQAIIIAIGTGIHDEFDIGKLRYHKI
VLMAIAIVSGHISI...LILLFRMRFLIENGVFLAQFLYIKKWQRSPEFAYSCR
FRCGLIEAGLKAGKKINKEDGIQRYKLGSEMDKKELWRI..MDPSVRVLRQVLDAAA
ADELFSILMGELVDAPRSFIRNAXDVRFLDV"

gene

7302..9318
/gene="gyrA"
/note="Rv0006"
/db_xref="EMBL:3261645"

CDS

/gene="recF"
/note="Rv0003", (MTCY10H4.01), recF gene, len: 385,
TR:E208598; contains PS00017 ATP/GTP-binding site motif A,
PS00617RecF protein signature 1, PS00616 RecF protein
signature 2"
/codon_start=1
/transl_table=11

CDS

/gene="Rv0004"
/note="Rv0004, (MTCY10H4.02), len: 187, almost identical
to ORF187 in recF region, MTCHISEP_6, E228999 :187 aa)
(97.9% identity in 187 aa overlap)"
/codon_start=1
/transl_table=11

gene

9914..10828
/gene="Rv0007"
/db_xref="EMBL:3261645"
9914..10928
/gene="Rv0007"
/note="Rv0007, (MTCY10H4.05), len: 304, Unknown, but
similar to 270722|MLCB1770_7 Mycobacterium leprae cosmid
B1770 (303 aa). FASTA scores: Opt: 812 z-score: 593.0 E():
1.6e-25, 54.2% identity in 319 aa overlap"

CDS

/gene="gyrA"
/note="Rv0006, (MTCY10H4.04), len: 838, gyrA, dna gyrase
subunit a, almost identical to GYRA_MYCTU Q07702, contains
PS00018EF-hand calcium-binding domain"
/codon_start=1
/transl_table=11
/product="gyrA"
/protein_id="CAB02427.1"
/db_xref="GI:1552559"
/db_xref="SWISS-PROT:Q07702"
/translation="MTDIILPPDDSLDRIEPVDIEQEMQRSYIDYAMSVIVGRALPEV
RDGLKPVHRRVLYAMFDSGFRPDRSHAKSARSVAEIMGNYHPHGDASLYDSLVRMAQP
WSLRYPLVDGQGNFGSSPGNDPPAAMRYTEARLTPLAMEMLREIDEETVDFIPNYDGRV
QEPTVLPSRFPNLLANGSGGIAVGMATNIPPHNLRELADAVFWALENHDADEEITLAA
VMGRVKGPDFPTAGLIVCSGQGTADAYKTCRGSIRHMGVEVEFDDSRGRKTSLVITELPY
QVNHDNFISIAEQVRDGK.AGISNIEDQSSDRVGLRIVIEIKRDAVAKVVINNLYKH
TQLQTSFGANMAIVDGVRTLRDQLLRYIVDHQLDVIVRRTIYRLRKANERAHILR
SKYXALADLDVIAIRASETVDIARAGLLELLDIEIQAQAI.DMQLRRLAAAPROR
IIDDLAK.EAEIADLEDI.AKPERQRGTVREFLAEIVDRHGDRRTRIIAADSCEE
DLIAREDVVVIIFGYYNRKTKIDLYRSQKRGGKGVQGAGLK.DDIVAHFFVSSTHDL
LIFFTIQGRVYRAKAYDLPEASRTARGQHVANLLAFQPFERIAQVI...QTGTDAYIV
LATRNG.VYKKSLTDPDSNRSGGIVAVNLRGNEDLVGAVICSACDDILLVSANGQSIR
FSATEAELRPMQRATSGVQGMRNICDRLISLNVVHEGTYL.VAISGYAKRTALEEY
PVQGRGKGVLIVMDRRGRSLVAALIVDDDSELTAVTSCGGVIHTAARQVRKAGROT
KGVFLMNLGEGUVTLLAIARNAEESGDDNAVDANGAQVTGN"

align_1

Qy   2701 GGGATCTTGGCCCA...TCATAGTTCTTGATCCTATCGGGGGGATCAGGAGGAGTGGCTTGCCA 2760
Db 1698715 GGGATCTTGCCCA...TCGATAGTTCTTGATCTATCGGGGATCAG AGGAGTGCCTTGCCA 1698774

Qy   2761 AAGGACACGTTCCGCGGGTCTCGAGAGCCCTCCGAACGAACTATTGAAACTTGATGATCGCG 2820
Db 1698775 AAGGACACGTTCCGCGGGTCTCCGAGAGCCCTCCGAACGAACTGATCAGAGTGATCGCG 1698834

Qy   2821 TCGGATCGGACTGTGGGATCGTCCCCGGGAAGCCTCGCTGGCCACAACAGAAAGGCTGGCATA 2880
Db 1698835 TGGGATCGACTGTGCGGGATCGTCCCCGGGAAGCCTCGTGGCCACAACAGAGAAAGGCTGGCATA 1698894

Qy   2941 AACACCACGCCGGAGAACCACGCCCGGGATCCGATCCACAACCCACGGCCGCGAAC 1698954
Db 1698895 TTGGACACACCACCACCCCGGGATCCGATCCACAACCCACGGCCGCGAAC 1698954

Qy   3001 GTGGCGGAGGCCCACCCGGTGATCCCCAATACCGGGCCCATTGTCGAGGCCGCAAAC 3060
Db 1699015 GTGGGGAGGCCCACCCGGTGCGGGACAACGGTAACCCACCGCGGATCGGTACC 1699074

Qy   3061 ACGAACCCGTGGCAGGTCCCTCGCGAATGGAGCCCCCACGCCGGCGCGGCGGGCCC 3240
Db 1699135 AGTGTCACGGAAGATGACCACTCGCCTCGCTCAGACCACACGTGCCCCCCCAGATGGCG 1699194

Qy   3121 AGTGTCACGAGAGATGGGCTGGGTAGTCGGAAAATCGCATACGGGTATGGCGGGGGGGGG 3180
Db 1699195 GGAGATATGACCGGCCATCGTTAGGGCGGCGACCAGGTAACACATCAGCCCCAAAGCGGCGCA 1699254

Qy   3181 CGGACATATGACCGGCCAATCGTTAGGGGCGGTAACACATCAGCCCCAAGCCCGGCGGCGCA 3240
Db 1699255 CCCAGCCAAATGGCTGGGTAGTCCGAAAATCGCATACGGGTATGCGGGGGGCCCCCCAAGCA 1699314

Qy   3301 GCAACGGGGGTCGGGCTGGTAATCGCGGCGGTACGGAGATCAGTAGTCCGGCGGAATGGGT 3360
Db 1699315 GCAACGGGGGTCGGGCTCGTTAATCGCGGGGTAGGGAGATCAGGAGTCGGCGGAATGGGT 1699374

Qy   3361 TGGAATCCCGGCGGCGGCCATGCCGTAGTCGTCGAAGCAGTGCCGACCGGATAGTCATG 3420
Db 1699375 TGAATCCCGGCGGCGGGACCGTAGTCGGCAGTAGGGCCGACCGATTACTG 1699434

Qy   3421 CCCCAATCAAGCCAACAGTCCCCCAGGGACTACCAAAAGAGTGCACCCG 3480
Db 1699435 CCCCAATCAAGCCAACACTCCCCCAGGGACTACCAAAAGAGTCGACCCG 1699494

Qy   3481 TACCACTGACTGAGGGGTGCGGCTCGGATTGGGAGGACGCGCCCGGCCCGGAAATCACCGGA 3540
Db 1699495 TACCACTGACTGAGGGGTCGGCTCGGATTGGGAGGACGCGCCCGGCCCGGAAATCACCGGA 1699554

Qy   354 CCGACCATGAGGGGTTGGGTCGCGTTCATCACTGTGAATGGTGG 3600
Db 1699555 CCGAGCAATGAGGGGTGCGGCTAGCGGTGGGTAGGCGGTCGGGTGCAGCGGGTGG 1699614

Qy   3601 GTCACGGCACCGCGCCCACAACACCTCCCGCGGGGTCGCTTCACTCTGAATTCTCTCCGGTTCTGCG 3660
Db 1699615 GTCACGGCACCGCGCCAACACCCGCGGGTCGCTTCACTCTGATTCTCTCCGGTTCTGCG 1699674

Qy   3661 CGAAAGGTAGCACAGAGCCCATCGTTCCCAACGCGGGGTCGCCTCTAGACCGGATC 3720
Db 1699675 CGAAACGTAGCAGAGAGCCCATGGTTGCAACGGCGGGTCGCCTCTAGACCGGATC 1699734

Qy   3721 TTCCTCGTTGGCAACCACAGGACGTCGTTGCCCAAAAGGGCCGCCTGGGCACCGAGATCT 3780
Db 1699735 TTCCTCGTTGGCAACGACAAGAGGACGTCGTTGCCCAAAGGGCGGCGGCTGGGCACCGAGATCT 1699794

Qy   3781 AGGATGAACCCACAGCCCGACGTTATGCCATGCGGAGAAGGGGCGACCGGGCAGGAGG 3840
Db 1699735 AGGATGAACCCACAGCCCGACGTTATGCCATGCGGAGAAGGGGCGACCGGGCAGGAGG 3840

Db 1699795 AGGATGAAGCCACAGCCAGACGAGCAATGCCGCGAGTATCGCATGGGGAACAGGAGG 1699854

Qy   3841 GGGAACCCCAGTGAAGGCACAGCGCGTA...AACGGAATCAGGAGGTCATCCATGGGAACAGGAGG 3900
Db 1699855 GGGAACCCAGTGAAGCGCACAGCGGTCAGCATACAGGACGGCGGCGGTCGACTTC 1699914

Qy   3901 GCTAAGCTCCCGCGTGAAAGGAATAGTGCGCGTCGTAGCCGCGGCGGAGCGTATCTGC 3960
Db 1699915 GCTAAGCTCCCGCTGAAAGGATATGTGGCCGCGGTTAGCCCGGACGAGGTCTATTCTGC 1699974

Qy   4021 AATCACGAATCACCGGTCAGGGCAGAGACATCATGATCGTAATCCTGACAGCATACGC 4080
Db 1699975 TGGCGGACACGGAATATCGGCAATGGCGGTAATGGGCCCTCAACAGCATCGTTC 1700034

Qy   4081 GTGCGGCGACCAGGCGTTGATGTGCAACGCTGATCGATCGAATTGACGGCCACCCAT 4140
Db 1700035 AATCACGAATCACCGGTCAGGGCAGAGACATCATGATCGTAATCCTGACAGCATACGC 1700094

Qy   4141 TATCACTATTGGGCGTTCCCGGCAGGTCATCGATACGGAGAGTGACGAATCGGTGT 4200
Db 1700095 GGTCGGCAAGAGGCGCGATCGATCGATACGGAGAGTGACGCGAATAGGTGT 1700154

Qy   4201 TCGGTTGCAGTAATAATTCGAACCGTGTGAACCGCATGCATGCAGAGATTCTTTCCCCAT 4260
Db 1700215 TCGGTTGAGTAATAATTCGAACCGTGTGAACCGCATGCATGCAGAGATTCTTTCCCCAT 1700274

Qy   4261 CCGGTCACGATGTCGATGCGGACGGCGCAGAATGGGTTGGTTACTATCGGATGCGGCGCGC 4320
Db 1700275 CCGGTCACGATGTCGATGCGGACGGCGCAGAATCGGTTCACTGTTGCATACGATGCGGCGCGC 1700334

Qy   4381 AGCATCGCTGGAAGTGGCGGGGGGTACTCGTGCGCAACTGGCGGCGGAGGCG 4440
Db 1700335 CGAGGCCGATTCGCACCGGACGGCGACTGCGGCCGTCGGTGCGTCCGCTGC 1700394

Qy   4441 ATTTGACACGCGCATGCAGCGGACGGCAAGGTTGGCGAAGTTCTGCATCGGATTG 4500
Db 1700395 AGGATCGCTGGAAGTGGCGGGGGTACTCGTGCGCAACTGGCGGCGGGAGCG 1700454

Qy   4501 CGACAGTCGTCCGTGAATTGGCGATTGGCGATTGTCCCAGACCCGCGCGGGGTTGTCGTGTGC 4560
Db 1700455 ATTTGACACGCGCATGCAGCGGACGGCAAGGTTGGCGAAGTTCTGCATCGGATTG 1700514

Qy   4561 CTTTATTCCGGAAANCGGTTCGACTCGAACTGCCGACCAACGTGCCGTTGTCGTGTGC 4620
Db 1700515 CGACAGTCGTCCGTGAATTGGCGATTGGCGATTGTCCCAGACCCGCGCGGGGTTGTCGTGTG 1700574

Qy   4621 CCTCGATCGGACGAGGTAGTCACAGGGGTTCGACTCGAACTGCCGACCAACGTGCCGTTGT 4680
Db 1700575 CTTTATTCCGGAAACGGTTCGACTCGAACTGCCGACCAACGTGCCGTTGTCGTGTG 1700634

Qy   4681 CACCAAAGAGGTGACATCAGCATGCAGGGAATCGATCGATACGGGCATGGA 4740
Db 1700635 CCTCGATCGGACGAGGTAGTCACAGGGGTTCGACTCGAACTGCCGACCAACGTGCCG 1700694

Qy   4741 CACCAAAGAGGTGACATCAGCATGCAGGGAATCGATCGATACGGGCATGGA 4800
Db 1700755 CACCAAAGAGGTGACATCAGCATGCAGGGAATCGATCGATACGGGCATGGA 1700814

Qy   4861 TCGGCGCGAAGGATGCACGGCTCACCTGGCTTCGCACCGTGCAAATGTTGGTATGGATGGT 4920
Db 1700815 TGACCGGCGAAGGGCAGAGCTGGGCGTTCGGCTCAACGGATTACCGGTTCCCAAAGGGAAAACGGCT 1700874

align_1

```
Qy   11461  TGTGCCATCACTTGGTTAGGCGCCTCGGGTTGGTCCTGGGTTGTTCGCGCAATGGTCAGGCTC  11520
            |||||  |||| ||||||||  |||||| ||||| |||||||||||||||||||||||||||||
Db 1707475  TGTGCCATCACTTGGTTAGGCCCCTCGGGTTGGTCCTGGGTTGTTCGCGCAATGGTCAGGCTC  1707534

Qy   11521  GACAGCCGCGTCGGAGCGCAGAGCGCACACATCCACGAGCGCCCCCTGGGGCTCTGGGCGTCG  11580
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1707535  GACACCCGCGTCGGAGCGCAGAGCGCACACATCCACGAGCGCCCCCTGGGGCTCTGGGCGTCG  1707594

Qy   11581  CGGCGGCGSAGCGTTCGCGCACTGGGTCTTGAGGCGCCGCGGCGGGGTGCGCCGGGGTGTT  11640
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1707595  CGGCGGCGSAGCGTTCGCGCACTGGGTCTTGAGGCGCCGCGGCGGGGTGCGCCGGGGTGTT  1707654

Qy   11641  GCAGCGGCAGGATGCCANTGGCGGATCGCGAGGATGGGTTGTTCCGTCCAGCGGCGGCGCC  11700
            || || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1707655  GCAGCGGCAGGATGGCGATCGCGGGGATGGGTTGTTCCGTCCAGCGGCGGCGCC  1707714

Qy   11701  CGTCCCGSGGGCCTGGAACGTTGGTCGTGTGGCCGAAGGATCACGCCGGCCATGACGGCGGCCGATGA  11760
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1707715  CCTCCCGSGGGCCTGGAACGTTGGTCGTGTGGCCGAAGGATCACGCCGGCCATGACGGCGGCCGATGA  1707774

Qy   11761  GTGCCTCGCCGCCGGGGGGTCTCCCGGGAACACGTTCGAGGCATCGCCGTCAAACGTTGCCGCAT  1820
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1707835  GCTCCTCGCCGCCGGGGGGGTCTCCCGGGAACACGTTCGAGGCATCGCCGTCAAACGTTGCCGCAT  1707894

Qy   11821  GCCCGCCGGACCCTGGTCAACACGAGAACTTGGCGGCTCCACCACCTCGGTTCTGTGCCCATGC  11880
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1707835  GCCCGCCGGACCCTGGTCAACACGAGAACTTGGCGGCTCCACCACCTCGGTTCTGTGCCCATGC  1707894

Qy   11941  CGTGGGCGTGGCGGACGTGCAGACAGCCAGGGTCGGCGAAGTGTGGCCCATGCA  11940
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1707895  CTTGGGCGTGGCGGACGATACACGGCCCATGGTGTCGGCGAAGTGTGGCCCATGCA  1707954

Qy   11941  CCGCGTSGCCGACGTCGAGAATACAACGAGGGCATGCAGGGGGCATGA.GTCGGCAGGAAATGTCTTCGTT  12000
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1707955  CCGCGTGGCCGACGTCGAGAATACAACGAGGGCATGCAGGGGGCATGA.GTCGGCAGGAAATGTCTTCGTT  1708014

Qy   12001  ACGGCGCTCGAGGCGACGGCTACACGGCCGAGCGATGCGATGGGTCGGCGTGAGTTCATCAAGAATCAACG  12060
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1708015  ACGGCGCTCGAGGCGACGGCTACACGGCCGAGCGATGCGATGGGTCGGCGTGAGTTCATCAAGAATCAACG  1708074

Qy   12061  CATCCCAGGCTTCAAGTTCCCGSCGCAACGGGAAGTTCTCTGGTCGTCATCGAAGTCATCAAGAGTCAACG  12120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1708075  CATCCCAGGCTTCAAGTTCCCGSCGCAACGGGAAGTTCTCTGGTCGTCATCGAAGTCATCAAGAGTCAACG  1708134

Qy   12121  CGTCCCCATAATCACTCGGACGCGGGATGGAAAGSTACCGGCGTCGTCGTCTTGGGCGGCGGTCGA  12180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1707955  CGGCCCCATAATCACTCGGACGCGGGATGGAAAGSTACCGGCGTCGTCGTCTTGGGCGGCGGTCGA  1708014

Qy   12241  ACGGGACGTA.TCGCCCGCACACCGCGCACGGCGGCGCCCGGCGTTCCCATTCGAGGTCGGA  12270
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1708025  ACGGGACGTACTCGCCCGAGCGCTGCACGGCGGCGCCCGGCGCGCAGTTCCCATTCGAGGTCGGA  1708314

Qy   12371  GAATGGACGGTAGCTCACGTTGGGGGCTACGTGGGCAGGCGCCCTCGGGGCACTGTCACACGGA  12420

Qy   12541  CCGGGCGCCAAGGGTATTGCGATGCGGCAAGGTCAAGGTGATGCCGTGGCCGTCGGCCGGCGTTGGGC  12600
```

RESULT 2

```
LOCUS       AE000516     4403836 bp    DNA    circular    BCT    27-APR-2001
DEFINITION  Mycobacterium tuberculosis CDC1551, complete genome.
ACCESSION   AE000516
VERSION     AE000516
KEYWORDS    .
SOURCE      Mycobacterium tuberculosis CDC1551.
ORGANISM    Mycobacterium tuberculosis CDC1551
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Mycobacteriaceae;
            Mycobacterium; Mycobacterium tuberculosis complex.
REFERENCE   1  (bases 1 to 4403836)
AUTHORS     Fleischmann,R.D., Alland,D., Eisen,J.A., Carpenter,L., White,O.,
            Peterson,J., DeBoy,R., Dodson,R., Gwinn,M., Haft,D., Hickey,E.,
            Kolonay,J.F., Nelson,W.C., Umayam,L.A., Ermolaeva,M.,
            Salzberg,S.L., Delcher,A., Utterback,T., Weidman,J., Khouri,H.,
            Gill,J., Mikula,A. and Bishai,W.
TITLE       Whole genome comparison of Mycobacterium tuberculosis clinical and
            laboratory strains
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 4403836)
AUTHORS     Fleischmann,R.D., Alland,D., Eisen,J.A., Carpenter,L., White,O.,
            Peterson,J., DeBoy,R., Dodson,R., Gwinn,M., Haft,D., Hickey,E.,
            Kolonay,J.F., Nelson,W.C., Umayam,L.A., Ermolaeva,M.,
            Salzberg,S.L., Delcher,A., Utterback,T., Weidman,J., Khouri,H.,
            Gill,J., Mikula,A. and Bishai,W.
TITLE       Direct Submission
JOURNAL     Submitted (25-APR-2001) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1..4403836
                     /organism="Mycobacterium tuberculosis CDC1551"
                     /strain="CDC1551"
                     /db_xref="taxon:83331"
                     /note="clinical strain"
     gene            1..1524
                     /gene="MT0001"
     CDS             1..1524
                     /gene="MT0001"
                     /note="similar to SP:P46388 PID:606325 PID:1262352;
                     identified by sequence similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="chromosomal replication initiator protein DnaA"
                     /protein_id="AAK44224.1"
                     /db_xref="GI:13879042"
                     /translation="MTDDPGSGFTIVWNAVVSELNGDPKVDDKSPSDANLSAELTPQQ
                     RAWLNLVQPLTIVEGFAL.SVPSSFVQNEIERHLRAPIIDALSRKLGHUTLGVRIAP
                     PAIEEADDTTVPPSENPAIISPDTTNDEIDDSAAARDGNDHSWPSYTTERPHNIDS
                     AIAGVTSLNRRYTFDTFVIGASNRFAHAAALAIAEAPARAYNT.FIWGESGLCKTHLL
                     HAAGNYAORLFPGMRVKYVSTEETNDFINSLRDDRKVAFKSSYRGVDVLLVDIOFI
                     ESKEGIOEFFNINTLHANKIVISSDRPPKQLATLEDRLRTRFEWGLITVQPPE
                     LERIAIRKKAQMERLAVPDDVLELASSIERNIRELEGALIRVTAFASLNKTPIDK
                     ALAEAVLRDLIADANTMQLSAAIIMAATAFYPRTVEVELERGPKKTRALGSRQIAMYL
                     CRELCLSLPKIGQAFGRDHTTVNYAQKKILSEMAEFREVFEHWKEJITKIRORSKR"
     gene            2052..3260
```